GAN-Tree: An Incrementally Learned Hierarchical Generative Framework for Multi-Modal Distributions

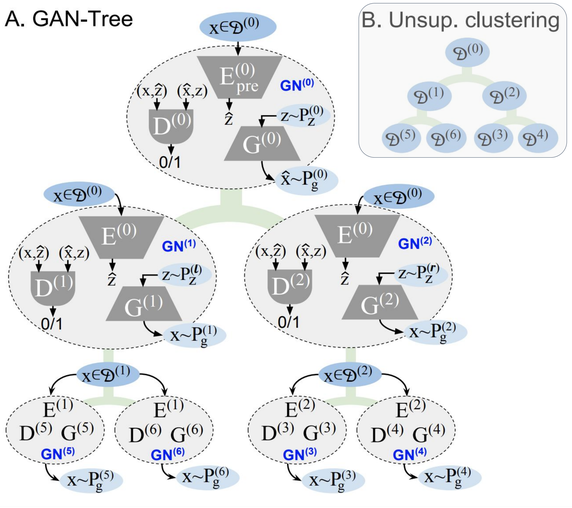
Jogendra Nath Kundu* Maharshi Gor* Dakshit Agrawal R. Venkatesh Babu

(*equal contribution) Video Analytics Lab, Indian Institute of Science, Bangalore, India

ICCV 2019 Seoul, Korea

Motivation

- GANs are limited in the following aspects:
  - Choice of latent variable setting (dimension, type of prior etc.)
  - Not great at mapping discontinuities (computational limitation)
- Previous proposed solutions for the above problem:
  - Single generator with multi-modal prior or Multi-generator n/w
- Limitations of available solutions:
  - Relies on an assumption of initial number of modes
  - Lack of flexibility for the user while generating new images
    - More modes → good Quality, but lacks Generalizability
    - Less modes→ lacks Quality but good Generalizability



Our Contributions

- We propose a novel generative modeling framework which:
  - allows flexible choice of the number of modes to be considered
  - allows flexibility in generation of new samples based on varied preference of quality versus diversity
  - allows incremental addition of new samples from similar but different classes introduced later in the absence of initial data samples

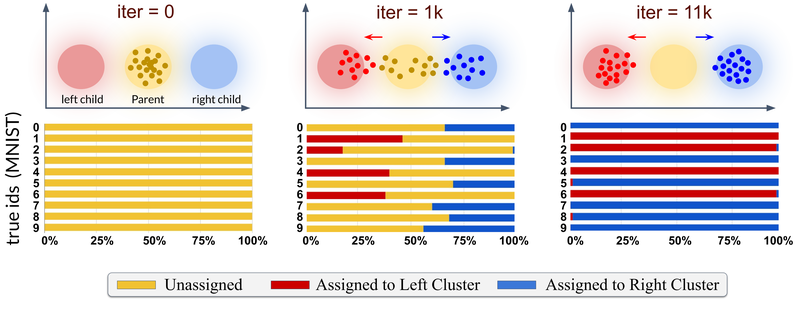
Training Algorithm

1: Starts with one GAN which is trained to model the whole dataset

2: Iterates over the following two algorithms until desired flexibility

3: Mode Split Procedure

- Top-down divisive clustering by leveraging the highly discriminative semantic characteristics in a fully-unsupervised manner

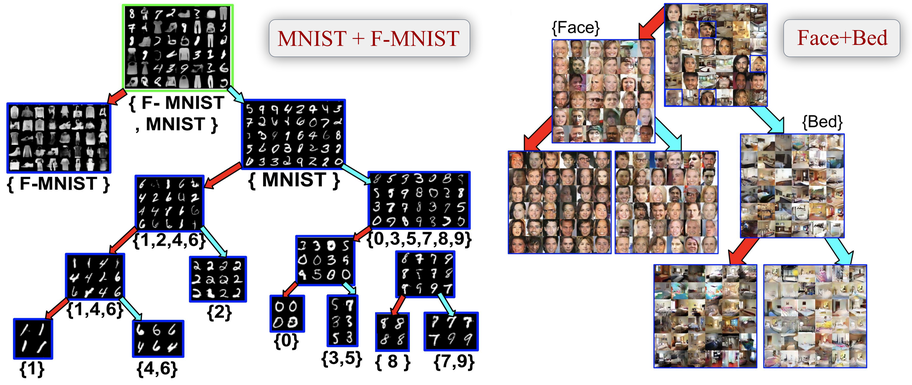


4: Bi-Modal GAN Training

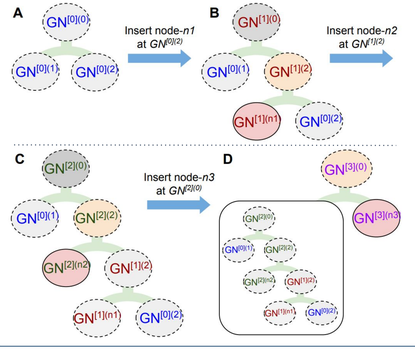
- ensure matching of the generated distribution with the expected target distribution (after unsupervised clustering of true images into two children groups) for faithful image generation

Experiments

- GAN-Tree generation for mixed datasets:



- Incremental GAN-Tree *iGAN-Tree*



- Quantitative comparison

Table 3: Inception (IS) and FID scores on CIFAR-10 and Imagenet dataset computed on 5K with varied number of generators.

Method	#Gen	CIFAR-10		ImageNet		
		IS ↑	FID ↓	IS ↑	FID ↓	#Param
GMVAE [9]	1	6.89	39.2	-	-	-
ClusterGAN [28]	1	7.02	37.1	-	-	-
RFGAN [3] (*root-node*)	1	6.87	38.0	20.01	46.4	50M
BigGAN (w/o label)	1	7.19	36.7	20.89	42.5	50M
MADGAN [13]	10	7.33	35.1	20.92	38.3	205M
DMWGAN-PL [21]	10	7.41	33.1	21.57	37.8	205M
Ours *GAN-Set* (ALI)	3	7.42	32.5	-	-	-
Ours *GAN-Set* (ALI)	5	7.63	28.2	-	-	-
Ours *GAN-Set* (RFGAN)	3	7.60	28.3	21.97	34.0	65M
Ours *GAN-Set* (RFGAN)	5	**7.91**	**27.8**	**24.84**	**29.4**	**105M**
Ours *GAN-Set* (BigGAN)	3	8.12	25.2	22.38	31.2	130M
Ours *GAN-Set* (BigGAN)	5	**8.60**	**21.9**	**25.93**	**27.1**	210M

Acknowledgements: This work was supported by a WIPRO PhD Fellowship (Jogendra) and a grant from ISRO, India. We also thank Google Inc for the travel grant support.